

Package ‘CKNNRLD’

July 13, 2026

Title Clustering-Based K-Nearest Neighbor Regression for Longitudinal Data

Version 0.2.2

Description Implements the 'CKNNRLD' algorithm (Clustering-Based K-Nearest Neighbor Regression for Longitudinal Data) for improving K-Nearest Neighbor ('KNN') regression on longitudinal data through cluster-based partitioning and localized prediction. Offers enhanced computational efficiency and accuracy for high-volume longitudinal datasets. The acronym 'KNN' stands for K-Nearest Neighbor. References: Loeloe MS, Tabatabaei SM, Sefidkar R, Mehrparvar AH, Jambarsang S (2025). ``Boosting K-nearest neighbor regression performance for longitudinal data through a novel learning approach." BMC Bioinformatics, 26, 232. <doi:10.1186/s12859-025-06205-1>.

License GPL-3

Encoding UTF-8

RoxygenNote 8.0.0

Imports Directional, graphics, grDevices, Rfast, stats

Depends R (>= 3.5.0)

NeedsCompilation no

Language en-US

Suggests knitr, rmarkdown, testthat

LazyData true

Author Mohammad Sadegh Loeloe [aut, cre],
Seyyed Mohammad Tabatabaei [aut],
Reyhane Sefidkar [aut],
Amir Houshang Mehrparvar [aut],
Sara Jambarsang [aut, ths]

Maintainer Mohammad Sadegh Loeloe <mslbiostat@gmail.com>

Repository CRAN

Date/Publication 2026-07-13 07:20:08 UTC

Contents

benchmark.CKNNRLD	2
BestC	3
CKNNRLD	4
CKNNRLD.tune	5
CKNNRLD_example	6
coef.CKNNRLD	7
fitted.CKNNRLD	7
importance.CKNNRLD	8
KNNRLD	9
KNNRLD.tune	10
performance.CKNNRLD	11
plot.CKNNRLD	11
predict.CKNNRLD	12
print.CKNNRLD	13
residuals.CKNNRLD	13
simulateCKNNRLD	14
summary.CKNNRLD	15
Index	16

benchmark.CKNNRLD	<i>Benchmark CKNNRLD vs KNNRLD</i>
-------------------	------------------------------------

Description

Compares CKNNRLD and KNNRLD on the same data.

Usage

benchmark.CKNNRLD(x, y, xnew, ytest, k = 5, c = 4)

Arguments

- | | |
|-------|------------------------------------|
| x | Matrix of predictors (training). |
| y | Matrix of responses (training). |
| xnew | Matrix of predictors (test). |
| ytest | Matrix of actual responses (test). |
| k | Number of neighbors. |
| c | Number of clusters. |

Value

A data frame with benchmark results.

Examples

```
data(CKNNRLD_example)
benchmark.CKNNRLD(
  x = CKNNRLD_example$x,
  y = CKNNRLD_example$y,
  xnew = CKNNRLD_example$x,
  ytest = CKNNRLD_example$y,
  k = 5, c = 3
)
```

BestC

*Find Optimal Number of Clusters for Longitudinal Data***Description**

This function determines the best number of clusters (C) for longitudinal data clustering using the elbow method (WCSS).

Usage

```
BestC(Y, range_clusters = 2:4, method = "kmeans")
```

Arguments

Y A matrix or data frame of longitudinal outcomes (subjects x timepoints).
range_clusters A numeric vector of cluster numbers to evaluate (e.g., 2:4).
method Clustering method to use (currently only "kmeans").

Value

A list with best_c, criteria, and criteria_best.

Examples

```
set.seed(123)
n <- 20
T <- 3
y <- matrix(rnorm(n * T), nrow = n)
best_c_info <- BestC(Y = y, range_clusters = 2:3)
print(best_c_info$best_c)
```

CKNNRLD

Clustering-based KNN Regression for Longitudinal Data (CKNNRLD)

Description

This function implements a clustering-based KNN regression method for longitudinal data.

Usage

```
CKNNRLD(
  x,
  y,
  xnew = NULL,
  k = 5,
  c = 4,
  cluster_method = "kmeans",
  return_model = FALSE
)
```

Arguments

<code>x</code>	Matrix of predictors (training set).
<code>y</code>	Matrix of longitudinal responses (training set).
<code>xnew</code>	Optional matrix of predictor values for test data. If <code>NULL</code> , uses <code>x</code> .
<code>k</code>	Number of nearest neighbors.
<code>c</code>	Number of clusters.
<code>cluster_method</code>	Clustering method (currently "kmeans").
<code>return_model</code>	Logical; if <code>TRUE</code> , returns a model object of class <code>CKNNRLD</code> .

Value

If `return_model = FALSE` (default), a data frame with predictions. If `return_model = TRUE`, a list of class `CKNNRLD` containing model info.

Examples

```
set.seed(123)
n <- 20; T <- 3; d <- 2
x <- matrix(runif(n * d), nrow = n)
y <- matrix(rnorm(n * T), nrow = n)
result <- CKNNRLD(x = x, y = y, k = 3, c = 2)
head(result)
```

Description

Automatically selects the best number of clusters (C) and tunes CKNNRLD.

Usage

```
CKNNRLD.tune(
  y,
  x,
  nfolds = 10,
  folds = NULL,
  seed = NULL,
  A = 10,
  C_range = 2:4,
  cluster_method = "kmeans"
)
```

Arguments

y	Matrix of longitudinal outcomes.
x	Matrix of predictor variables.
nfolds	Number of folds for cross-validation.
folds	Optional list of pre-specified fold indices.
seed	Random seed for reproducibility.
A	Maximum number of neighbors to evaluate.
C_range	Range of cluster numbers to evaluate.
cluster_method	Clustering method to use (currently only "kmeans").

Value

A list containing best_c, cluster_results, cluster_sizes, etc.

Examples

```
set.seed(123)
n <- 20
T <- 3
d <- 2
x <- matrix(runif(n * d), nrow = n)
y <- matrix(rnorm(n * T), nrow = n)
tune_result <- CKNNRLD.tune(
  y = y,
  x = x,
```

```

    nfolds = 3,
    A = 4,
    C_range = 2:3
  )
  print(tune_result$best_c)

```

CKNNRLD_example

Simulated longitudinal data for CKNNRLD example

Description

A synthetic dataset containing longitudinal responses and predictors for demonstration of the CKNNRLD package.

Usage

```
CKNNRLD_example
```

Format

A list with two elements:

- x** Matrix of 100 observations and 3 predictors
- y** Matrix of 100 observations and 5 time points
- clusters** True cluster assignments
- cluster_centers** True cluster centers
- parameters** Simulation parameters

Details

The data was generated using `simulateCKNNRLD(n = 100, T = 5, d = 3, C = 3)` with seed 123. It contains three distinct trajectory patterns.

Examples

```

data(CKNNRLD_example)
str(CKNNRLD_example)
head(CKNNRLD_example$x)
head(CKNNRLD_example$y)

```

coef.CKNNRLD	<i>Extract model information from CKNNRLD</i>
--------------	---

Description

Extract model information from CKNNRLD

Usage

```
## S3 method for class 'CKNNRLD'  
coef(object, ...)
```

Arguments

object	A fitted CKNNRLD model object.
...	Additional arguments.

Value

A list containing model information.

fitted.CKNNRLD	<i>Extract fitted values from CKNNRLD model</i>
----------------	---

Description

Extract fitted values from CKNNRLD model

Usage

```
## S3 method for class 'CKNNRLD'  
fitted(object, ...)
```

Arguments

object	A fitted CKNNRLD model object.
...	Additional arguments.

Value

A matrix of fitted values.

importance.CKNNRLD	<i>Variable Importance for CKNNRLD using Permutation</i>
--------------------	--

Description

Computes permutation-based variable importance for CKNNRLD models.

Usage

```
importance.CKNNRLD(
  x,
  y,
  xnew,
  ytest,
  k = 5,
  c = 4,
  nperm = 10,
  metric = "MSE",
  plot = TRUE
)
```

Arguments

x	Matrix of predictors (training set).
y	Matrix of longitudinal responses (training set).
xnew	Matrix of predictors for test data.
ytest	Matrix of actual responses for test data.
k	Number of nearest neighbors.
c	Number of clusters.
nperm	Number of permutations (default: 10).
metric	Performance metric: "MSE", "RMSE", or "MAE" (default: "MSE").
plot	Logical; if TRUE, plots variable importance (default: TRUE).

Value

An object of class "CKNNRLD_importance".

Examples

```
data(CKNNRLD_example)
imp <- importance.CKNNRLD(
  x = CKNNRLD_example$x,
  y = CKNNRLD_example$y,
  xnew = CKNNRLD_example$x,
  ytest = CKNNRLD_example$y,
```

```

    k = 5, c = 3, nperm = 5
  )
  print(imp)

```

KNNRLD

Standard K-Nearest Neighbor Regression for Longitudinal Data

Description

This function performs KNN regression for longitudinal data without clustering. It predicts longitudinal outcomes for new observations based on the average of their k nearest neighbors in the predictor space.

Usage

```
KNNRLD(xnew, y, x, k = 5)
```

Arguments

<code>xnew</code>	A matrix of predictor values for prediction (test set).
<code>y</code>	A matrix or data frame of longitudinal responses (training set).
<code>x</code>	A matrix or data frame of training predictor values.
<code>k</code>	Number of nearest neighbors to use. Can be a scalar or a vector.

Value

A list of matrices with predicted values for each value of k . Each matrix has dimensions `nrow(xnew)` x `ncol(y)`.

Examples

```

set.seed(123)
n <- 20
T <- 3
d <- 2
x <- matrix(runif(n * d), nrow = n)
y <- matrix(rnorm(n * T), nrow = n)
train_idx <- sample(1:n, 14)
test_idx <- setdiff(1:n, train_idx)
pred <- KNNRLD(
  xnew = x[test_idx, ],
  y = y[train_idx, ],
  x = x[train_idx, ],
  k = 3
)
head(pred[[1]])

```

KNNRLD.tune

Tune k in KNNRLD using Cross-Validation

Description

Finds the optimal number of neighbors for KNN regression using k-fold CV.

Usage

```
KNNRLD.tune(
  y,
  x,
  nfolds = 10,
  folds = NULL,
  seed = NULL,
  A = 10,
  graph = FALSE
)
```

Arguments

y	Matrix of longitudinal outcomes.
x	Matrix of predictor variables.
nfolds	Number of cross-validation folds.
folds	Optional list of pre-specified fold indices.
seed	Optional random seed.
A	Maximum number of neighbors to evaluate.
graph	Logical; if TRUE, plots MSPE vs. k.

Value

A list containing crit, best_k, performance, and runtime.

Examples

```
set.seed(123)
n <- 20
T <- 3
d <- 2
x <- matrix(runif(n * d), nrow = n)
y <- matrix(rnorm(n * T), nrow = n)
tune_result <- KNNRLD.tune(
  y = y,
  x = x,
  nfolds = 3,
  A = 4
)
```

```
)
str(tune_result)
```

performance.CKNNRLD	<i>Performance metrics for CKNNRLD models</i>
---------------------	---

Description

Performance metrics for CKNNRLD models

Usage

```
performance.CKNNRLD(object, ytest = NULL, ...)
```

Arguments

object	A fitted CKNNRLD model object.
ytest	Optional test set actual responses.
...	Additional arguments.

Value

A list of performance metrics.

plot.CKNNRLD	<i>Plot method for CKNNRLD objects</i>
--------------	--

Description

Provides visualization for CKNNRLD results.

Usage

```
## S3 method for class 'CKNNRLD'
plot(x, y_actual = NULL, type = "clusters", ...)
```

Arguments

x	An object returned by CKNNRLD or a fitted CKNNRLD model.
y_actual	Optional matrix of actual longitudinal responses.
type	Type of plot: "clusters" (default), "centers", or "comparison".
...	Additional arguments passed to plot().

Value

A plot (invisibly returns NULL).

Examples

```
data(CKNNRLD_example)
fit <- CKNNRLD(x = CKNNRLD_example$x, y = CKNNRLD_example$y, k = 5, c = 3,
               return_model = TRUE)
plot(fit, type = "clusters")
```

predict.CKNNRLD	<i>Predict method for CKNNRLD objects</i>
-----------------	---

Description

Predicts longitudinal responses for new data using a fitted CKNNRLD model.

Usage

```
## S3 method for class 'CKNNRLD'
predict(object, newdata = NULL, ...)
```

Arguments

object	A fitted CKNNRLD model object (output from CKNNRLD with return_model = TRUE).
newdata	Optional matrix of new predictor values.
...	Additional arguments passed to CKNNRLD.

Value

A matrix of predicted longitudinal responses.

Examples

```
data(CKNNRLD_example)
fit <- CKNNRLD(x = CKNNRLD_example$x, y = CKNNRLD_example$y, k = 5, c = 3,
               return_model = TRUE)
pred <- predict(fit, newdata = CKNNRLD_example$x[1:5, ])
head(pred)
```

print.CKNNRLD	<i>Print method for CKNNRLD objects</i>
---------------	---

Description

Prints a concise summary of a fitted CKNNRLD model.

Usage

```
## S3 method for class 'CKNNRLD'  
print(x, ...)
```

Arguments

x	A fitted CKNNRLD model object.
...	Additional arguments passed to print.

Value

Invisibly returns the object.

Examples

```
data(CKNNRLD_example)  
fit <- CKNNRLD(x = CKNNRLD_example$x, y = CKNNRLD_example$y, k = 5, c = 3,  
               return_model = TRUE)  
print(fit)
```

residuals.CKNNRLD	<i>Extract residuals from CKNNRLD model</i>
-------------------	---

Description

Extract residuals from CKNNRLD model

Usage

```
## S3 method for class 'CKNNRLD'  
residuals(object, ...)
```

Arguments

object	A fitted CKNNRLD model object.
...	Additional arguments.

Value

A matrix of residuals.

simulateCKNNRLD

Simulate Longitudinal Data for CKNNRLD

Description

Generates synthetic longitudinal data for testing and demonstration.

Usage

```
simulateCKNNRLD(
  n = 100,
  T = 5,
  d = 2,
  C = 3,
  noise_sd = 1,
  random_sd = 1,
  seed = NULL
)
```

Arguments

n	Number of subjects (default: 100).
T	Number of time points (default: 5).
d	Number of predictors (default: 2).
C	Number of clusters (default: 3).
noise_sd	Standard deviation of noise (default: 1).
random_sd	Standard deviation of random intercept (default: 1).
seed	Optional random seed.

Value

A list containing x, y, clusters, cluster_centers, and parameters.

Examples

```
sim_data <- simulateCKNNRLD(n = 50, T = 4, d = 2, C = 3)
str(sim_data)
```

summary.CKNNRLD	<i>Summary method for CKNNRLD objects</i>
-----------------	---

Description

Provides a detailed summary of a fitted CKNNRLD model.

Usage

```
## S3 method for class 'CKNNRLD'  
summary(object, ytest = NULL, ...)
```

Arguments

object	A fitted CKNNRLD model object.
ytest	Optional test set actual responses.
...	Additional arguments.

Value

A list containing summary statistics.

Examples

```
data(CKNNRLD_example)  
fit <- CKNNRLD(x = CKNNRLD_example$x, y = CKNNRLD_example$y, k = 5, c = 3,  
               return_model = TRUE)  
summary(fit)
```

Index

* **datasets**

CKNNRLD_example, [6](#)

benchmark.CKNNRLD, [2](#)

BestC, [3](#)

CKNNRLD, [4](#)

CKNNRLD.tune, [5](#)

CKNNRLD_example, [6](#)

coef.CKNNRLD, [7](#)

fitted.CKNNRLD, [7](#)

importance.CKNNRLD, [8](#)

KNNRLD, [9](#)

KNNRLD.tune, [10](#)

performance.CKNNRLD, [11](#)

plot.CKNNRLD, [11](#)

predict.CKNNRLD, [12](#)

print.CKNNRLD, [13](#)

residuals.CKNNRLD, [13](#)

simulateCKNNRLD, [14](#)

summary.CKNNRLD, [15](#)